Online ISSN: 2456-1495
CODEN (USA): JSLOBJ

JOURNAL OF SCIENTIFIC LETTERS
www.jslsci.com

ENHANCED EEG SIGNAL PROCESSING FOR INTELLIGENT EPILEPTIC SEIZURE DETECTION

Hude Tejaswi Kiran

Research Scholar, Department of Computer Science Engineering, Vikrant University, Gwalior, M.P

Dr. Sanmati Kumar Jain

Research Supervisor, Department of Computer Science Engineering, Vikrant University, Gwalior, M.P

ABSTRACT

The goal of this study is to use two feature extraction approaches, Wavelet Packet decomposition (WPD) and Genetic Algorithm-Based Frequency-Domain Feature Search (GAFDS), to automate the extraction of electroencephalogram (EEG) data without referring to clinicians. In order to identify epileptic seizures, three machine learning methods were used: Random Forest (RF), Support Vector Machine (SVM), and Conventional Neural Networks (CNNs). With CNNs as the classifier and GAFDS as the feature extraction, the classifier results indicate a better accuracy rate of 98.21% compared to the SVM's 96.87% and RF's 89.16%.

Keywords: Quality, Machine Learning, Electroencephalography, Seizure Detection

Online ISSN: 2456-1495
CODEN (USA): JSLOBJ

I. INTRODUCTION

Seizures that happen repeatedly and without apparent cause are the hallmark of epilepsy, a long-term neurological disorder caused by an overactive and abnormal brain electrical activity. The condition has a profound impact on individuals's physical and mental health, as well as their educational opportunities, occupational status, and quality of life in general; it may strike at any age and affect millions of people globally. From short bursts of unconsciousness and odd motions to full-blown convulsions, seizures may take many forms. Thus, accurate and early seizure detection is crucial for epilepsy diagnosis, clinical management, and treatment timeliness. One of the most used non-invasive methods for monitoring brain activity and detecting aberrant electrical patterns linked to epileptic seizures is electroencephalography (EEG).

Electrodes affixed to the scalp capture electrical activity emanating from the brain, which is represented by an electroencephalogram (EEG) signal. The patterns included in these signals may vary between normal and epileptic states, and they provide important information on the brain's functional status. Seizures characterised by epilepsy are characterised by changes in the amplitude, frequency, rhythm, and temporal features of electroencephalogram (EEG) recordings. On the other hand, electroencephalogram (EEG) signals are intricate, dynamic, and prone to artefacts and noise brought about by things like eye movements, muscle activity, electrode displacement, and environmental interference. This makes neurologists' manual processing of lengthy EEG records a laborious and difficult task. There is a growing demand for trustworthy computational algorithms that can detect seizures efficiently and accurately, especially with the availability of continuous EEG monitoring on the rise.

An essential step in automated EEG seizure detection systems is feature extraction. The main goal of feature extraction is to separate non-seizure brain activity from raw EEG recordings by creating a collection of relevant and discriminative properties. Appropriate feature extraction techniques may simplify the data while retaining the most significant information needed for classification, which is especially helpful for high-dimensional, redundant raw EEG data. Nonlinear, time-frequency-domain, frequency-domain, and time-domain features are the basic classifications for the retrieved features. Statistical measurements including signal energy, skewness, kurtosis, standard deviation, variance, and root mean square are all part of the time-domain characteristics. The temporal behaviour and amplitude distribution of EEG signals can be inferred from these characteristics. Commonly derived from methods like the Fast Fourier Transform (FFT), frequency-domain characteristics characterise the signal energy distribution throughout several frequency bands, such as theta, beta, alpha, and gamma rhythms.

Due to the non-stationary nature of EEG data, time-frequency analysis has emerged as a crucial

E-mail: editor.jslsci888@gmail.com

method for seizure identification. Methods like the Discrete Wavelet Transform (DWT), Wavelet Transform (WT), and Short-Time Fourier Transform (STFT) allow for the simultaneous study of EEG data in the time and frequency domains. Seizure occurrences may induce brief and localised alterations in EEG activity, which wavelet-based feature extraction can record. Since epileptic EEG signals may display variations in complexity and dynamical behaviour, nonlinear analytic approaches have also garnered a lot of interest. Additional information about the underlying characteristics of brain activity can be found using features such as fractal dimension, entropy, approximate entropy, sample entropy, permutation entropy, and Lyapunov-related measures. The capacity of a detection system to differentiate between seizure activity and normal or interictal EEG patterns can be enhanced by combining various types of features.

In order to create automated seizure detection systems based on electroencephalograms, machine learning algorithms are crucial. Machine learning algorithms may discover patterns from annotated EEG data and categorise recordings into distinct states, including seizure and non-seizure, once pertinent information have been retrieved. There has been a lot of research on traditional supervised learning algorithms that have been used for this purpose, such as Support Vector Machines (SVM), Artificial Neural Networks (ANN), Decision Trees, Random Forests, and Naïve Bayes classifiers. The capacities of these approaches to deal with complex feature distributions, nonlinear relationships, and high-dimensional data vary. A seizure detection system's efficacy is contingent upon the relevance and quality of the characteristics that are retrieved, as well as the machine learning technique that is used.

Recent years have seen a surge in interest in EEG signal analysis as a potential application of deep learning and other machine learning techniques. In order to automatically learn complicated representations from EEG data, algorithms like CNNs, RNNs, LSTM networks, and hybrid deep learning architectures may be used. On the other hand, when working with small datasets or when concerns about computational efficiency and interpretability arise, traditional feature extraction can still be useful. To enhance classification performance and remove unnecessary or redundant information, dimensionality reduction and feature selection approaches may be used.

Automatic seizure detection systems may be built upon a promising architecture that combines EEG signal processing, feature extraction, and machine learning. Reduced manual EEG interpretation, continuous monitoring, and faster seizure event detection are all ways in which an efficient system might aid neurologists. Still, there are obstacles that impact system effectiveness, such as changes in recording settings, class imbalance, noisy EEG recordings, and inter-patient variability. Research into the creation of reliable feature extraction methods that are also compatible with appropriate machine learning algorithms is, therefore, crucial. The development of more effective automated seizure detection systems may be advanced by a thorough examination of EEG characteristics and how well they perform in machine learning-based categorisation. This, in turn, can lead to better diagnosis of epilepsy and better treatment for patients.

II. REVIEW OF LITERATURE

Hossain, Fahima et al., (2023) In this study, we provide a method that uses Machine Learning techniques to forecast epileptic seizures using electroencephalogram (EEG) data, with the goal of pharmacologically preventing seizures. These electrical impulses in the brain are often detected using electrocorticography (ECoG) and electroencephalography (EEG) devices. In addition to producing a mountain of data, these signals are complex, noisy, non-linear, and non-stationary. Seizure detection and understanding the brain's operations are thus challenging tasks. Machine learning classifiers can identify seizures, categorise EEG data, and highlight relevant patterns without compromising performance. The dataset including information on epileptic seizures was categorised using several classifiers in this research. The results showed that support vector machines outperformed several other classifiers, including: Decision Tree, Stochastic Gradient Descent (SGD), Naive Bayes, K-Nearest Neighbours, Random Forest, Logistic Regression, Bagging, AdaBoost, Gradient Boosting, Multi-layer Perceptron (MLP), and XGBoost. We put our proposed technique to the test on 22 subjects using the CHBMIT dataset of scalp EEG signals in this research. Our proposed method for seizure prediction achieves impressive results: an accuracy rate of 95.88%, a recall rate of 86.91%, a precision of 1%, and a sensitivity of 1%.

Al-Huseiny, Muayed&Sajit, Ahmed. (2023) Using a publicly available dataset that included EEG recordings from both healthy individuals and those with epilepsy, this study developed three low-complexity classifiers: a decision tree, a random forest, and an AdaBoost algorithm. Short electrical waves reflecting brain activity were first extracted from the data during preprocessing. Once the models have been chosen, the signals are fed into them. Based on the experimental findings, decision tree had the second-highest accuracy rate (96.93%) in detecting epileptic seizures in EEG data, whereas random forest had the greatest overall accuracy rate (97.23%). With an accuracy of 87.23%, AdaBoost was the algorithm with the worst performance. Decision tree, random forest, and AdaBoost all achieved AUC ratings of 99%, 99.9%, and 95.6%, respectively. Even though state-of-the-art classifiers have a larger temporal complexity, our findings are on par with them.

Mouleeshuwarapprabu, R. &Kasthuri, N.. (2023) The diagnosis of epileptic seizures is often accomplished with the use of electroencephalography (EEG). Visual examination of EEG recordings over an extended period of time, however, is characterised by its inherent subjectivity, lengthy process, and propensity for mistake. To address these issues, several seizure detection algorithms are used; nonetheless, it is worth noting that these approaches exhibit a significant ratio of false positives. This study thus presents a new method for automated seizure identification with the goal of decreasing the percentage of false positives. According to the suggested strategies, users should pick the one that works best for their specific classification issue. A signal's temporal evolution across different time scales can be described using information about local regularity, which is the basis of the wavelet transform. It is necessary to build a signal that is suitable for the data type in order to extract Multi-Model Features from separate data sources. Feature extraction from an ECG signal is distinct from feature extraction from any other source. Because it may speed up the seizure detection process and alleviate visual analysis overload, the suggested approach for automated epilepsy identification is a promising alternative to the current state of the art.

Brari, Zayneb&Belghith, Safya. (2021) Seizures, a symptom of epilepsy, a debilitating neurological disorder, impact persons' daily life. Since encephalography is easy to acquire, inexpensive, and produces substantial findings, it is the principal test for epilepsy diagnosis. However, a neurological consultation is necessary for the exploitation of electroencephalogram (EEG) information; this takes a lot of time and isn't always an option. A number of studies have recently suggested various machine learning algorithms for the categorisation of electroencephalogram (EEG) data in order to identify the three distinct brain states of healthy, seizure-free periods, and seizure-prone people. In addition to ensuring accurate programs, developed systems must have minimal storage memory requirements, run as quickly as possible, and handle real-time application requirements. A basic model with a small number of easily computed and classified features should therefore form the basis of any proposed systems. Here, we present a fresh strategy for extracting EEG features, one that relies on a different way to find the Correlation Dimension (CD). Our solution outperforms previously published research in the experimental test conducted using a benchmark database. We attain a 100% accuracy rate for almost all classification issues using specified subset combinations, which is a reliable result. We use a fast-running model with few features to keep the program simple.

Mardini, Wail et al., (2020) A state-of-the-art electroencephalogram (EEG) may provide important details about a variety of brain disorders. The epileptic seizure is one such anomaly. We provide a system that can distinguish between normal and epileptic EEG data, and use that information to identify seizures. The goal of the proposed method is to automatically distinguish between normal and abnormal signals. The goal of this effort is to decrease computing expenses and increase the accuracy of detecting epileptic seizures. This is addressed by the proposed framework, which employs the 54-DWT mother wavelets analysis of EEG signals with the Genetic algorithm (GA) in conjunction with four other ML classifiers: Naive Bayes (NB), Artificial Neural Network (ANN), Support Vector Machine (SVM), and K-Nearest Neighbours (KNN). We test the efficacy of these four ML classifiers on fourteen distinct permutations of two-class epilepsy detection. The four classifiers outperformed each other and produced similar results for the generated statistical features from the 54-DWT mother wavelets, according to the experimental findings. However, the ANN classifier beat the others and attained the highest accuracy in most dataset combinations.

III. RESEARCH METHODOLOGY

Figure 1 shows the raw data generated from the EEG dataset, which is the starting point for the suggested technique of this work. In order to get the dataset ready for analysis, we shall remove any irrelevant or noisy data. Once that is done, the data may be prepared for the feature extraction techniques (WPD, GAFDS) to begin the preprocessing step. The PCA feature selection method will then be applied to the pre-processed data. Lastly, the features that were chosen will be fed into the ML algorithms that were suggested (SVM, CNN), and then the outcomes will be assessed and contrasted. We will thoroughly examine each of the steps outlined in the proposed methodology in the sections that follow.

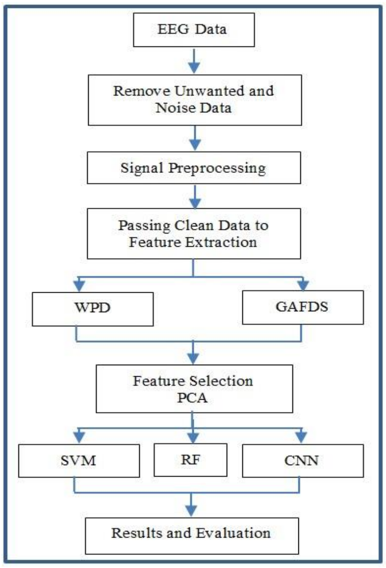

Figure 1. The proposed methodology

Dataset Description

Recordings of electroencephalograms (EEGs) taken from children at the Children's Hospital Boston who suffer from uncontrollable seizures have been compiled. The 22 subjects included 5 males ranging in age from 3 to 22 and 17 females ranging in age from 1.5 to 19. The recordings were then organised into 23 cases. There are nine to forty continuous.edf files for a single subject in each case of the data files. Although some cases'.edf files are two or four hours long, the majority of them contain one hour of digitalised EEG signals. On occasion, files that document seizures tend to be shorter. A sampling rate of 256 s/s and a resolution of 16 bits were used for all signals. With a few exceptions, the majority of files include twenty-three EEG signals. A comprehensive list of all 664.edf files in this collection can be found in the dataset's RECORED file. Among them, 129 files have one or more seizures listed in the RECORDS-WITHSEIZURES file. There are a total of 198 seizures documented here, up from 182 in the initial collection of 23 instances.

Feature Extraction

The EEG dataset will be processed using two feature extraction methods, namely WPD and GAFDS, in this study.

Machine Learning Algorithms

Using a training set of EEG feature sets that comprise occurrences with established class relationships, ML's EEG signal classification deals with the task of classifying a set of classes to which a new reading belongs. The data will be classified in order to create test segments using data acquired from several classifiers. Because of this, it can probably tell if the data is related to seizures or not. It is the intention of this work to employ SVM, RF, and CNNs. Anaconda Python 3.7 was utilised for the experimental work's software setup.

IV. RESULT AND DISCUSSION

The suggested approach makes use of and contrasts various ML techniques. This experiment makes use of an Intel i7 CPU and the Python programming language. One set of data was used for training purposes (70%) and the other set for testing purposes (30%). Based on the data shown in Table 1 and Figure 1, it can be concluded that CNN outperformed SVM for both feature extraction techniques. When compared to the other ML algorithms used for feature extraction, RF's accuracy was the lowest.

Table 1: Comprehensive Results

Model	Feature Extraction	Accuracy %	Precision %	Recall %	F1-Score %
SVM	GAFDS	96.87	97.42	98.31	97.86
	WPD	95.74	96.35	97.28	96.81
RF	GAFDS	89.16	88.74	97.63	92.99
	WPD	87.82	87.56	96.42	91.78
CNN	GAFDS	98.21	97.86	99.73	98.79

E-mail: editor.jslsci888@gmail.com

	WPD	97.48	97.12	99.18	98.14

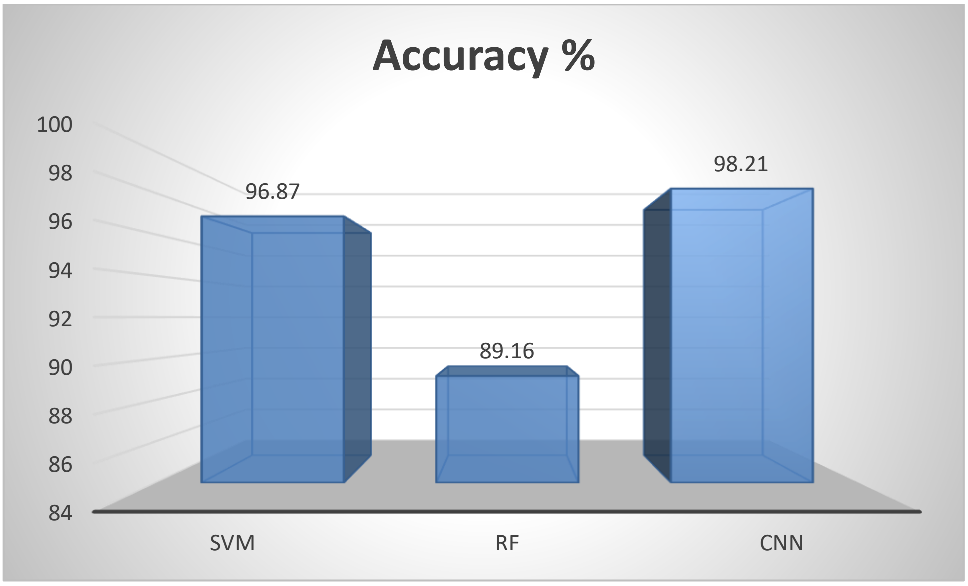

Figure 1: Accuracy results for 70% training and 30% testing of EEG signals.

There is a wide range of outcomes, from very low to very high accuracy, when it comes to identifying EEG data for datasets of varying sizes using various ML algorithms. Causes for these discrepancies include, but are not limited to, variations in EEG data screening and collecting procedures, feature selection, data formatting, and classification methods and parameters. Most of the time, the accuracy of EEG data categorisation is poor. The main cause of this is the difficulty in analysing EEG signals due to the large amount of noise and artefacts that are present in them. As mentioned before, there are specific techniques that may be used to eliminate artefacts from EEG data. Nevertheless, this might lead to the loss of important data and impact the accuracy of the categorisation findings.

V. CONCLUSION

In this work, the primary characteristics from the EEG signals were extracted using two feature extraction approaches and other pre-processing methods. Prior to the classification step, PCA was employed to determine which characteristics were most beneficial. The primary objective of this research was to evaluate several ML methods for EEG data categorisation. In the trials, three ML algorithms—specifically, SVM, RF, and CNN—were used. Data from electroencephalograms (EEGs) collected from children's hospitals in Boston, including reports from children with uncontrollable seizures, were used to evaluate the selected procedures. After developing the recommended algorithms, we evaluated their outputs using several metrics for quality assurance, and ultimately settled on the most effective classifier for EEG signal diagnosis.

Online ISSN: 2456-1495
CODEN (USA): JSLOBJ

REFERENCES

[1] F. Hossain, M. A. A. Walid, R. Akter, S. M. Galib, M. Mir, and Azad, "Early prediction of epileptic seizure using EEG spectral features and machine learning approaches," *Journal of Jilin University (Engineering and Technology Edition)*, vol. 42, no. 5, pp. 113–122, 2023.

[2] M. Al-Huseiny and A. Sajit, "Detection of epileptic seizures in EEG by using machine learning techniques," *Diagnostyka*, vol. 24, no. 1, pp. 1–7, 2023.

[3] D. Raab, A. Theissler, and M. Spiliopoulou, "XAI4EEG: Spectral and spatio-temporal explanation of deep learning-based seizure detection in EEG time series," *Neural Computing and Applications*, vol. 35, pp. 10051–10068, 2023.

[4] R. Mouleeshuwarapprabu and N. Kasthuri, "Feature extraction and classification of EEG signal using multilayer perceptron," *Journal of Electrical Engineering & Technology*, vol. 18, no. 4, pp. 3171–3178, 2023.

[5] Z. Brari and S. Belghith, "A novel machine learning approach for epilepsy diagnosis using EEG signals based on correlation dimension," *IFAC-PapersOnLine*, vol. 54, no. 17, pp. 7–11, 2021.

[6] K. Rasheed, A. Qayyum, J. Qadir, S. Sivathamboo, P. Kwan, L. Kuhlmann, T. O'Brien, and A. Razi, "Machine learning for predicting epileptic seizures using EEG signals: A review," *IEEE Reviews in Biomedical Engineering*, vol. 14, no. 3, pp. 139–155, 2021.

[7] Z. Brari and S. Belghith, "A novel machine learning approach for epilepsy diagnosis using EEG signals based on correlation dimension," *IFAC-PapersOnLine*, vol. 54, no. 17, pp. 7–11, 2021.

[8] W. Mardini, M. B. Yassein, R. Alrawashdeh, S. Aljawarneh, Y. Khamayseh, and O. Meqdadi, "Enhanced detection of epileptic seizure using EEG signals in combination with machine learning classifiers," *IEEE Access*, vol. 8, no. 2, pp. 24046–24055, 2020.

[9] M. Moridani, Z. Mahabadi, and N. Javadi, "Heart rate variability features for different stress classification," *BratislavskéLekárskeListy*, vol. 121, no. 9, pp. 619–627, 2020.

[10] B. Hosseinifard, M. H. Moradi, and R. Rostami, "Classifying depression patients and normal subjects using machine learning techniques and nonlinear features from EEG signal," *Computer Methods and Programs in Biomedicine*, vol. 109, no. 3, pp. 339–345, 2013.

E-mail:* *editor.jslsci888@gmail.com

Online ISSN: 2456-1495
CODEN (USA): JSLOBJ

[11] S. Zurek, P. Guzik, S. Pawlak, M. Kosmider, and J. Piskorski, "On the relation between correlation dimension, approximate entropy and sample entropy parameters, and a fast algorithm for their calculation," *Physica A: Statistical Mechanics and its Applications*, vol. 391, no. 24, pp. 6601–6610, 2012.

[12] R. Palaniappan and K. V. R. Ravi, "Improving visual evoked potential feature classification for person recognition using PCA and normalization," *Pattern Recognition Letters*, vol. 27, no. 7, pp. 726–733, 2006.

[13] S. Marcel and J. del R. Millán, "Person authentication using brainwaves (EEG) and maximum a posteriori model adaptation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 12, pp. 1107–1146, 2005.

[14] S. N. Sarbadhikari and K. Chakrabarty, "Chaos in the brain: A short review alluding to epilepsy, depression, exercise and lateralization," *Medical Engineering & Physics*, vol. 23, no. 7, pp. 447–457, 2001.